R E P O R T R E S U M E S

ED 019 674 24 AL 001 270

A DEVELOPMENTAL STUDY OF THE EFFECT OF RELATIONSHIPS BETWEEN WORDS ON A MEMORY TASK.

BY- FELZEN, ENID

REPORT NUMBER BR-5-0602 PUB DATE FEB 68
CONTRACT OEC-1-6-050602-0431
EDRS PRICE MF-$0.50 HC-$2.76 67P.

DESCRIPTORS- *ASSOCIATION (PSYCHOLOGICAL), PSYCHOLOGICAL STUDIES, *PSYCHOLINGUISTICS, LANGUAGE RESEARCH, *WORD RECOGNITION, *DEVELOPMENTAL PSYCHOLOGY, CHILD DEVELOPMENT, CHILDREN, GRADE 3, GRADE 6, ASSOCIATIVE LEARNING,

THIS STUDY WAS CONDUCTED TO EXPLORE DEVELOPMENTALLY THE FEATURES OF WORDS USED IN THE PROCESS OF ORGANIZING RELATED WORDS TOGETHER. A CONTINUOUS RECOGNITION MEMORY TASK OF 121 WORDS WAS USED IN WHICH SOME OF THE WORDS (EXPERIMENTAL WORDS) PRESENTED ONLY ONCE IN THE LIST WERE SEMANTICALLY OR PHONETICALLY RELATED TO OTHER LIST WORDS. EIGHTY THIRD- AND SIXTH-GRADE CHILDREN HAD TO INDICATE WHETHER THE WORDS HAD BEEN PRESENTED BEFORE IN THE LIST OR NOT. IN PARTICULAR, THIS THESIS WAS CONCERNED WITH FALSE RECOGNITION ERRORS WHERE A WORD WAS RECOGNIZED AS HAVING BEEN PRESENTED PREVIOUSLY. THE ERROR SCORE WAS SUPPLEMENTED BY A MEASURE OF THE REACTION TIME FOR RESPONDING. IT WAS FOUND THAT--(1) THE CHILDREN WERE MUCH MORE LIKELY TO RECOGNIZE A WORD CORRECTLY THAN TO MAKE AN ERROR, (2) IT TOOK SIGNIFICANTLY LONGER TO MAKE AN ERROR THAN TO RESPOND CORRECTLY, AND (3) THIRD-GRADERS MADE SIGNIFICANTLY MORE ERRORS TO RHYME WORDS THAN TO SEMANTICALLY RELATED WORDS WHILE THE SIXTH-GRADE CHILDREN MADE MORE ERRORS TO THE SEMANTIC THAN TO THE RHYME WORDS. THE AUTHOR FEELS THAT THESE RESULTS (1) MAY INDICATE THAT WORDS MAINTAIN THEIR SEPARATE IDENTITY AND CAN THUS BE CORRECTLY IDENTIFIED, (2) SUGGEST A PROCESS WHEREBY RELATED WORDS ARE GROUPED TOGETHER SO THAT AT THE PRESENTATION OF AN EXPERIMENTAL WORD THE CHILD MIGHT LINK IT WITH THE WORD TO WHICH IT IS RELATED AND WHICH WAS HEARD PREVIOUSLY IN THE LIST, AND (3) SUGGEST THAT DIFFERENT RELATIONSHIPS BETWEEN WORDS ARE USED FOR THE GROUPING OF WORDS AT DIFFERENT AGE LEVELS. THIS DOCUMENT COMPRISES A MASTER'S THESIS PRESENTED TO THE FACULTY OF CORNELL UNIVERSITY. (SEE RELATED DOCUMENT AL 000 833.) (DO)

BR-5-0602
OEC-1-6-05062-0431
PA 24

ED019674

A DEVELOPMENTAL STUDY OF THE EFFECT OF RELATIONSHIPS BETWEEN WORDS ON A MEMORY TASK

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University for the Degree of

Master of Arts

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

by

Enid Felzen

February, 1968

ERIC
TL 001 270

INTRODUCTION

Psychologists have been concerned with the psychological processes of encoding and storing of words in memory; processes which make words available for later use. More specifically, there has been interest in specifying the characteristics or features that two or more words may share so that some form of organization is facilitated between the words. As Anisfeld and Knapp (1967) have pointed out, linguists have emphasized that vocabulary is not a list of unrelated items but rather has structure based on the "relations of words to categories or to other words." Psychologists have used different methodological approaches to investigate what these organizational processes might be.

The methods range from conditioning techniques to tasks of learning and recall. One approach to be used in the present study is that of continuous recognition memory. The S is presented with a list of words, where some of the words are related to words presented earlier in the list. The S's task is to indicate whether each word had been read before in the list or not. In particular, this method is used to study the errors in recall produced by words related to earlier presented words. These are referred to as intrusion errors or errors of false recognition. An analysis of such errors may shed light on the way in which the S organizes related words.

The adult and developmental research on the organization of words will be reported in the following pages. There are only a limited number of developmental studies in this area, but qualitative age differences have been reported, namely that the younger child is more apt to make use of phonetic characteristics in relating words while the older child uses semantic relationships (relations of synonymity and antonymy). The evidence of such differences reveals additional dimensions used in the organization of words beyond those found from studies limited to adults. The finding that young children are more responsive to physical similarities of words than to semantic relations contributes additional information for the construction and elaboration of a theory of language development. In view of the significance of a developmental study in this area of word relations and because of some of the methodological weaknesses in the few existing studies, a re-evaluation of the age differences seems warranted. This study will employ the continuous recognition memory technique to investigate this age differential.

LITERATURE REVIEW

Research with Adult Ss

The conditioned generalization design has been the most widely used method to study the effect of semantic and phonetic relationships of words. Essentially the paradigm is one in which a word, the conditioned stimulus, is paired with a response, the conditioned response, through a process of reinforcement, referred to as the conditioning phase of the experiment. Then the synonym, antonym and homonym of the original word are tested for the presence of the conditioned response in the generalization phase of the experiment. Studies of semantic generalization have produced fairly consistent results to demonstrate first, that the phenomenon of generalization exists and secondly, that for adult Ss generalization is stronger to semantically related words than to phonetically related ones. Since Feather (1965) presents an extensive and intensive review of the research on semantic generalization, only a few selected studies have been included here.

The research studies vary as to the conditioned response measure employed, but the findings generally support the hypotheses. After conditioning the salivation response to four visually presented words, Razarn (1939) found that in comparison to the originally conditioned stimulus, Ss salivated 59% to the synonyms and only 37% to the homonyms of these words. These findings were replicated by Riess (1940) using

the galvanic skin response as the conditioned response. For synonyms of the conditioned words the GSR increased 141 % over the pretraining session, while to homonyms the increase was only 94.5%. Similar results have been reported with the conditioned heart rate (Lacey and Smith, 1954) and with the eye lid response (Hartman, 1963). Thus, the majority of studies do strongly suggest that semantic relations based on synonyms or antonyms and to a much lesser extent common phonetic characteristics, represented by homonyms are used by adults in the organization of words.

Feather (1965) in critically reviewing the methodological designs of 25 semantic generalization studies, reports that despite the relative consistency of the results, many of the research designs do not provide adequate controls against the possibility of pseudoconditioning or do not present sufficient evidence of generalization. As a consequence of the reinforcement procedure there occurs an increased response sensitization to all stimulation evidenced in the S's increased base rate in responding to all stimuli. Therefore, the possibility exists that in studies without adequate control stimuli in the initial phase or in the generalization phase of the experiment that the increase in amplitude response to the conditioned stimulus or the demonstrated transfer of the conditioned response to the generalization stimuli may be artifacts of the S's increased sensitization to all stimuli. Thus, such results do not unequivocally demonstrate that either conditioning or generalization had occurred.

Feather's criticism does not reduce the importance of the generalization differences reported for the synonyms, antonyms and homonyms. The generalization scores may be inflated because of the S's increased tendency to respond, but the relative differences still remain.

Most of the experiments of semantic generalization seemed to be concerned with the conditioned response used to measure generalization and not with the specific characteristics of the generalization stimuli. The stimuli were characterized as either homonym, antonym or synonym. Another method, that of clustering in free recall has investigated more qualitative features of semantic relations. Besides synonymy and antonymy relations, the variables of association and ease of categorization were evaluated for their role in organization. The organization of words which are related to each other because they belong to the same category (e.g., apple, orange, pear) was compared to that of words not categorically linked (e.g., cat, table, father); the organization of words commonly associated with each other as measured by word association norms (e.g., house-home) was compared with that of words infrequently associated (e.g., thief-crook).

The method of clustering in free recall developed from the observation that when Ss were asked to list items according to a category, they produced words in sequential order such that words sharing some essential characteristic within the category group would follow each other (Bousfield, 1953). For example in the bird category, Ss might produce two separate clusters like: "hawk, eagle, vulture" and then

perhaps "chicken, turkey, duck." A cluster is thus defined by two or more words that can be identified as having similar characteristics. From these observations, Bousfield postulated that the production of items into clusters "is a consequence of organization in thinking and recall" and that the testing for recall of related items could reveal information about this underlying organization.

The method of clustering in free recall is one in which the S is presented with a list of either categorically or associatively linked words. The words are given in random order so that related words do not appear together and the S is then required to recall the words in "any order." The results of clustering studies indicate that randomly presented associated words are recalled in sequential units of related words.

Initially, clustering experiments emphasized the influence of the categorical relations, but gradually the emphasis shifted to associative relations. Bousfield (1953), using list items from four different categories, found that Ss recalled the words in terms of the superordinate categories represented. However, further studies found that categorical relations were insufficient for facilitating clustering in recall. Bousfield, Cohen, Whitmarsh (1958) presented Ss with words having high taxonomic frequency, i.e., words given frequently as associates to a category name, and low taxonmic words, i.e., words not frequently given as associates to a category name. The results indicated that high taxonomic words produced more clustering in recall than

the low taxonomic ones. Thus, the more "readily" a group of words can be categorized in a class, the more "readily" they will be recalled and the greater will be the organization of the recall.

It has also been shown that words which are commonly associated but not necessarily linked by categorical relations, will yield clustering in recall (Jenkins et al.,1958). Further support for the role of associative relations in the organization of words is presented by Cofer (1959). He found that synonyms yield only minimal clustering when the words are not also highly associated with each other. This coupled with the differential amount of clustering as a function of taxonomic frequency, shifted the emphasis from categorical to associative relations in an explanation of the clustering process. Furthermore, these studies place greater emphasis on the role of associations in the process by which words are organized than would have been suggested by the studies of semantic and phonetic generalization.

The method of false recognition memory, originally introduced by Shepard and Teghtoosian (1961), permits one to infer the way in which Ss organize related words by the intrusion errors the related words produce on the memory task. Underwood (1965) presented college Ss with a continuous list of 200 words at the rate of one word per 10 seconds. The S was to decide whether each word had been presented earlier in the list or whether it had not been played before in the list. The design called for some words to be presented only once in the list but to be associates of words that were given before in the list. A measure of false recognition was based on the S's errors in

indicating that a word had occurred earlier when in fact it was the first presentation of the word. Errors of failing to recognize a word as having occurred before also exist but were not of concern in this study.

Underwood's list was constructed of essentially three different types of words: (a) Critical stimulus words: words whose associate was presented later in the list. (b) Experimental words: the most frequently produced associates to the "critical stimulus" words in word association tests. These words always followed the words to which they were associatively related, with some words appearing in the intervening positions. (c) Control words: other common associates whose stimulus counterpart was not present in the list and which had no relationship to other list words.

The study was restricted to words where the relationship between the critical stimulus and its experimental word was one of association. More specifically, there were four categories of associates: (a) opposites; (b) superordinates where the experimental word represented the name of a category class of which the critical stimulus belonged (e.g., sparrow-bird); (a) converging associates where the experimental word was an associate of the stimulus word sharing a contiguous relationships (e.g., bed-sleep); (d) sense impressions (e.g., doughnut-round).

The results, except for the words representing sensory impressions, confirm the prediction that Ss will falsely recognize having heard a word before in the list if the word is associatively related to a word

that did occur earlier in the list. The Ss made significantly more errors to experimental words than to control words. Mean number of errors to experimental words = 4.43 while mean to control words = 2.53 ($t = 8.26$, $p < .001$). Another important result was that Ss were more likely to correctly detect the word as not having been presented earlier in the list than in falsely recognizing a non-repeated word.

Underwood's theoretical explanation of the occurrence of errors is based on the mediational properties of "implicit associative responses" (IAR). "The particular IAR to a given word is assumed to be the most frequent associate produced to the word in word association procedures." Underwood hypothesized that this IAR is made to the actual stimulus word and therefore when this highly associated response is later presented to the S he becomes confused and responds as if he had heard it earlier.

Anisfeld and Knapp (1967) using the same technique, investigated the role of snyonymy as well as association in yielding errors on this memory task. It was reasoned that "the constant use of paraphrasing in everyday life communication suggests that in coding for memory under normal conditions speakers retain mostly the semantic content of a message." Even though the nature of this memory task emphasizes verbatim coding as evidenced by the small percentage of false recognition errors, it was hypothesized that since synonyms are often substituted for each other in normal speech they should be responsible for some of the intrusion errors on this task.

Therefore, the experimental word was either a synonym, an associate or both a synonym and associate of the "preceding" related word.

Besides introducing the relationship of synonymity, Anisfeld and Knapp delineated more specific criteria for the control words. Each control was matched to its experimental word on the basis of frequency of occurrence as measured by the Thorndike-Lorge G. Count (1944); part of speech and syllablic length. The purpose of these criteria was to control for the effect of factors other than semantic relations in the production of false recognition errors.

The results of this study were consistent with the hypothesis that more errors would be made to associates and synonyms than to controls. (Mean number of errors to synonyms = 1.46; mean for controls = .71, $t = 2.94$, $p < .01$. Mean number of errors to associates = 1.89; mean for controls = 1.00, $t = 2.52$, $p < .02$). It should be noted that in this study even more than in Underwood's, the Ss were much more likely to identify the words correctly than to make an error in recognition.

To explain the occurrence of false recognition errors, Anisfeld and Knapp introduced a feature model where the encoding of a word is seen as the "simultaneous activation" of semantic, syntatic and physical features. The word is thus stored as a "complex of features." "Thus when a new word is heard which shares some significant features with an old word, the S may be led mistakenly to 'disregard' the distinguishing features and consider the two as identical."

Briefly then, the studies on word relations thus far reported indicate some of the semantic characteristics used by adults in the organization of words. The findings of semantic generalization studies differentiated between the effect of phonetic and meaningful relations. Essentially the results indicated that adults generalize more to synonyms and antonyms than to homonyms. A qualitative investigation of the semantic relationships stems from the studies of clustering in recall and false recognition memory. Generally, these studies report that associational, categorical and synonymous dimensions are effective in the organization of words.

Developmental Studies of Word Relations

The differential effect that semantic and phonetic relations have on the organization of words has not been widely studied developmentally. Most often cited is a study by Riess (1946). Using a conditioning paradigm with four different age groups, he was able to demonstrate a developmental gradient of generalization with respect to synonyms, antonyms, and homonyms.

There were five conditioning sessions for each subject and in any one session all the conditioned stimuli and generalization words were presented but only one conditioned stimulus was reinforced. In the conditioning period a buzzer was sounded throughout the exposure of the particular stimulus to be conditioned during that session. Conditioning was discontinued when the CS tested without the buzzer yielded gain in the electro-dermal response (EDR) at least three times that of the preconditioning phase. Since no neutral stimuli were

presented at the time the conditioned stimuli were tested for gain in EDR over a preconditioning level, Feather's (1965) criticism is applicable that the gain might be a function of the S's increased base rate of responding to all stimulation.

Riess reported evidence of semantic generalization as well as a pattern of generalization reflecting age differences in the use of different verbal relations. For the mean 7 year, 9 month old age group, the greatest transfer was to homonyms. Percentage gain to homonyms was 71.58% which was significantly more than to antonyms or synonyms, 62.79% and 58.43% respectively. In the 10 year, 8 month age group, antonyms account for most of the transfer, 42.90% compared to 28.76% for synonyms and 26.41% for homonyms. However, in the 14.0 year olds while the least transfer is to homonyms, 24.69%, the effect of the semantically related words is reversed with significantly more generalization to synonyms, 45.22%, than to antonyms, 37.71%. The same pattern is repeated for the college group with most transfer to synonyms, 52.85% and least to homonyms, 18.56%. The results suggest a specific developmental gradient of generalization with a decrease in the transfer to homonyms with increasing age. Moreover, in the two younger groups Ss generalized least to the synonyms whereas in the older groups the synonyms are in first position with the antonyms in second place.

In addition to Riess' lack of control for pseudo-conditioning, discussed above, he provided no assurance that generalization actually occurred. Since no control words were presented at the time generalization measures were taken to the synonyms, antonyms, and homonyms, the

reported gain in EDR to these words may represent the S's increased response sensitivity to all stimulation. Finally the EDR's tendency to easy elicitation by any novel stimulus makes it a poor choice as a response measure in a conditioning study (Feather, 1965). Therefore, the absolute percentages may not be reliable but the relative differences are important.

Despite the methodological flaws in the Riess study, it does suggest developmental differences in the reaction to words of common semantic and phonetic characteristics. Riess concluded that semantic generalization reflects the individual's mode and level of language development. Riess further implied that the younger Ss may not have learned to make use of the meaningful relationships that exist between words but rather rely on similar physical characteristics in the organization of words. This suggests that the younger child at this level of language development may have verbal labels which are not well integrated into a semantic system. Such observations of developmental differences are in line with the research and theoretical suggestions on language development and cognitive growth. On the tasks of grouping objects, the younger child in the 7-8 year old age range, groups on the basis of physical features (size, color); not until 11-12 years does functional grouping become a dominant mode, where the child responds to the common use of the objects (Olver, 1961). The younger child is more likely to associate objects on concrete rather than conceptual grounds (Siegel, 1964).

Related to the research on the developmental differences in semantic generalization is Luria's work (1963) on the relationship between levels of retardation and differences in the orienting response. The procedure is one of presenting the S with a "signal word" to which he is instructed to respond with a button press. New words are then presented to the S and measurement is made for the presence of the vascular response* which the S previously emitted along with the button press to the "signal word." This vascular response emitted without the overt response to a new stimulus is defined as the orienting response.

Like the generalization technique, this permits an investigation of transfer of response from word to word based on the relationships that exist between the words. In Luria's terms, "an investigation of connections that exist among words making them act as if they belonged to families or groups." While the young child may not be able to verbalize the relationship that exists between the words, the emission of the vascular response indicates that for the S some connection existed between the original signal word and the new word.

Luria was able to show that Ss of different intellectual levels vary in their response to the different relationships connecting the words. A normal child of "junior school age" (he does not specify

---

* Vascular response or orienting response is defined as an observed constriction of the blood vessels of the hand, a galvanic skin response, a change in respiration, cardiac activity and an increase in muscle tone manifested at the presentation of a novel stimulus. Luria measured the vascular response as a constriction of the blood vessels in the fingers.

ages) did not respond to a word similar in sound to the signal word but only to words similar in meaning - both antonyms and synonyms of the signal word. However, a mildly retarded child reacted to words both similar in meaning and sound to the signal word. With a severely retarded child the orienting response was elicited by words similar in sound, and only by those semantic words of synonymous relations, (i.e., kitten-cat; but not dog - cat). These findings are supported by results of word association tests where the young child responds to the stimulus usually with a synonym but not with an antonym (Woodworth, 1938, p. 346). No mention is made of control words or of the extent of the Ss' familiarity with the words.

Luria presents no statistical data and cites only specific cases to illustrate his points. Furthermore, his conclusions are based on one S per cell -- One normal, one mildly retarded and one imbicile S. Even though there are experimental inadequacies, this study supports in part the pattern suggested by Riess' (1946) developmental study. With increasing age as well as increasing intellectual ability, the sound characteristics of words contribute less than the semantic features in the organization of words. However, from these results the relationship of synonyms and antonyms is unclear. In the Riess study the younger child transferred more to the antonyms than to the synonyms but in this study the synonyms but not the antonyms elicited the vascular response for the severely retarded child.

Rice (1963), in his doctoral dissertation later published by Rice and DiVesta (1965), attempting to avoid some of the procedural diffi-

culties encountered in the conditioning generalization studies introduced a series of learning tasks to measure developmental differences in generalization. The experiment was divided into three stages. First, there was a paired-associate learning task, similar to the conditioning phase of a generalization experiment where the S learned to pair the appropriate nonsense syllable with its stimulus word. After learning of the pairs had been established, effects of generalization were tested. The homonyms, antonyms, synonyms and control words for each training word were presented only once and generalization was measured by the S's ability to indicate the appropriate nonsense syllable for each test stimulus. The Ss were told specifically to "think of the words you have already learned. They will help you in saying the right answer."

A second assessment of generalization was made in the third stage with a savings in learning task. The same Ss were presented with the homonyms, antonyms, and synonyms of the original stimuli and the corresponding nonsense syllables. Generalization was measured by the reduced number of learning trials required as compared to the learning of the training word-nonsense syllable pairs.

The results of this study show that the semantic relations represented by synonyms, and antonyms become increasingly important in accounting for generalization with increasing age. On both generalization tasks, college students gave significantly more "correct" responses to antonyms and synonyms than did third graders. However, across ages there were no significant differences in the amount of

"correct" responses to homonyms. Furthermore, college students generalized equally to homonyms, antonyms and synonyms.

It was only in the college group that the pattern of results between the three types of test words remained consistent on both generalization tasks. The third and fifth graders on the first generalization test transferred more to homonyms than to synonyms or antonyms. However, on the savings in learning task the third graders did not respond significantly better to the homonyms than to the antonyms or synonyms. Moreover, the superiority of homonyms over semantic characteristics for the fifth graders evidenced on the first generalization task did not hold on the second where instead both phonetic and semantic relations contributed equally to savings in learning.

Besides the internal inconsistency of the results in the Rice study, the findings are not congruent with those of Luria (1963) or Riess, (1946). Both these latter investigators report a decreasing shift in the effectiveness of phonetic similarity for producing generalization with increasing age or with increasing intelligence when comparing the retarded to the normal range of intelligence. However, in the Rice study with increasing age semantic similarity became an effective dimension for generalization while the effect of homonyms was constant across ages. Rice attributes the stable effect of the homonyms to his task instructions where Ss were encouraged to pay attention to all cues of the words so that even for college students the phonetic characteristics were relevant dimensions. Furthermore, it should be noted that unlike Riess and Luria, Rice found no

differences in the generalization to antonyms and synonyms. Without further investigation it may be assumed that task differences between Rice's two generalization tests and between his tasks and the conditioning procedures of Riess and Luria are responsible for the different findings obtained.

There are certain procedural aspects of the Rice study which deserve further discussion. First of all he carefully controlled for the S's familiarity with the meaning of the words in order to demonstrate that the superiority of phonetic generalization does not represent the Ss's unfamiliarity with the words but rather a phase in development. Therefore, only those words whose synonyms, antonyms, and homonyms familiar to a group of third, fifth and seventh graders were used as stimuli. Secondly, the sample was selected from those children who demonstrated familiarity with the words.

Summary

To reiterate briefly, the results of studies with adult Ss report that adults generalize more to words representing semantic relations than to words of common phonetic characteristics. Furthermore, it will be recalled that highly associated words lead to more clustering in recall and to more errors on a continuous recognition task that infrequently associated words. Despite this emphasis placed on the dimension of association, Anisfeld and Knapp (1967) concluded that synonyms which were not common associates did cause errors on the recognition task.

The developmental studies present evidence for a pattern of generalization reflecting age differences. The young child (and severely retarded) generalizes more to phonetically similar words than to words semantically related. From the studies it is not clear if the effectiveness of the phonetic characteristics decreases with age. Riess reports that college Ss generalize more to the semantically related words than to homonyms, but Rice found no age difference in the generalization to homonyms. Moreover, the developmental differences with respect to synonyms and antonyms are not consistent. Reiss found that the two younger groups (7 year, 9 month olds and the 10 year, 8 month olds) gave more transfer of the conditioned response to antonyms than to synonyms, while the 14.0 year olds and the college Ss generalized more to synonyms than to antonyms. However, Luria found that the severely retarded child emitted the orienting response to synonyms only while the normal and mildly retarded child gave the response to both synonyms and antonyms. Finally, Rice found that the 10 year, 8 month old age group generalized more to antonyms than to synonyms but that the college Ss generalized to both synonyms and antonyms. Thus while the developmental difference for semantic characteristics per se is maintained across varied research designs, no clearcut pattern is upheld for differences between synonyms and antonyms, or for that matter, between the phonetic and semantic features of words.

In the present study, a continuous recognition memory task will be used where some of the words presented only once in the list will be semantically or phonetically related to other list words. The third and sixth grade Ss will have to indicate for each word whether

it had been presented before in the list or not. Based on the findings of adult and developmental studies on the organization of words, the following hypotheses were made about the Ss' performance on the task of recognition memory:

1. Ss would make more recognition errors to experimental words which are phonetically or semantically related to other words than to control words not related to any other words on the list.

2. The third grader will make more errors to phonetically related words than to semantically related ones. From the developmental studies it is not clear whether sixth graders will make more errors to the semantic words than to the phonetically related ones.

3. Within the errors made to semantic words, words representing common associates should account for more recognition errors on the memory task than words having a low association value, although the uncommon associates should contribute to errors on the task.

METHODOLOGICAL CONSIDERATIONS FOR THE PRESENT STUDY

Choice of Subjects

To demonstrate a developmental difference it was decided to select two ages which would represent the outer limits of the age difference reported by other investigators. Third graders were used as the younger age group because with this group both Riess and Rice and DiVesta report evidence of phonetic generalization. The older group represented a compromise. Because the performance of the fifth graders was not clearly differentiated in the Rice and DiVesta study, i.e., on one task their performance was similar to the third graders and on the second task it was more like the older groups' performance, in this study sixth graders were selected as the older group of Ss.

Choice of Method

Due to the methodological gaps of the developmental studies in this area, the present study was undertaken to investigate further the age differences with respect to semantic and phonetic similarity of words. A technique different from that used in other developmental studies was selected to confirm the results already reported. The proposed method was one of false recognition memory used by Underwood, and Anisfeld and Knapp. With this particular method the effect of semantic and phonetic relations among words is evaluated by the amount of confusion errors made by the words which are related in meaning or

sound to a particular word presented earlier to the S. The technique permits the measurement of two different types of errors: 1) False recognition errors where the S indicates that he heard the word earlier in the list when in fact it is the word's first presentation. This is the only type of error that can occur to words semantically or phonetically related to list words and to their respective control words since all of these words appear only once in the list. False recognition can also occur to the first presentation of any other list word. 2) Negative errors occur in repeated list words where the S fails to identify having heard the words before.

The procedure called for an aural presentation of the stimuli. In the previous studies of semantic and phonetic generalization, words were visually presented to the S on a screen. Since homonyms share certain graphmic features it is possible that visually presented homonyms produce generalization because of physical similarities other than sound (for example, similarity of the initial letter, or of other letters in the word). Thus aural presentation of words might facilitate the evaluating of the role of sound similarity of words, although the possibility was not eliminated that Ss might still act on the basis of other relationships.

Choice of Measures

The error measure of the amount of intrusion that semantic and phonetic related words cause on a memory task was supplemented by a measure of the reaction time for responses. The false recognition task maximizes the S's uncertainty as to whether a word preceded by

its synonym or associate had been heard just a few minutes earlier or not. Reaction time differences to words preceded by a related word and to control words unrelated to other words on the list might add information to aid in the making of inferences about the processes going on in the coding of words.

Choice of Stimuli

A further investigation of the effect of the synonymic and antonymic word relations seems warranted because of the inconsistent findings. Rice failed to find any developmental gradient in the amount of generalization to these different types of word categories. In Riess' analysis of the generalization to semantically related words the two younger groups transferred more to antonyms than to synonyms while the older groups generalized more to synonyms than to antonyms. However, Luria found that the severely retarded child will emit the orienting response to synonyms but not antonyms. Association value of the words was included in the study as an independent variable in an attempt to establish the developmental effect of synonymy and antonymy. The results of the clustering experiments (Bousfield, Cohen and Whitmarsh, 1958; Jenkins et al.(1958) as well as Underwood's study(1965) suggest that common associates of words contribute to the formation of clusters, improvement of recall, and the production of recognition errors on the false recognition task. Furthermore Cofer (1959) has shown that low associated synonyms yield only minimal clustering. Therefore, this study investigated the differential effect of synonyms and antonyms representing common and uncommon associates.

Because of the nature of the present design where stimuli are presented auditorily, rhyme words were used rather than homonyms to represent the phonetic relations of words (for example: "kitten-mitten"). It should be noted that the same words were used as stimuli at the different age levels. This introduced the problem that the older Ss might have been more familiar with the words than the younger Ss. However, at the same time it did eliminate the possibility that any reported developmental differences might be a function of different stimuli used at the different ages. A second grade vocabulary list (Walters and Courtis, 1948) was consulted for selection of all words to insure that the younger Ss would know the meanings of the words. This procedure did not control for differential familiarity with the words between the two ages.

Procedural Considerations

The stimuli were composed of three types of words: 1) target words; 2) experimental words; 3) control words. After having heard the target word the S had to make a decision regarding words related to it (experimental words) and unrelated words (control words). Each target word was presented only twice in the list, always before the appearance of its synonym, antonym or rhyme word. This was in contrast to the three presentations of the word used by Underwood (1965) and by Anisfeld and Knapp (1967) in one study. It was decided to reduce the number of presentations in order to shorten the task time, considering the younger ages of the Ss used in this study. While Underwood reported no false recognition errors when the "critical

stimulus word" was presented only once in the list before its associate, Anisfeld and Knapp in their second study found that two presentations of these words was sufficient to produce errors.

Each word in the list was presented only once in immediate succession because in pretest trials the immediate repetition of each word seemed to make the task boring without adding to a clearer perception of the word. Furthermore, word repetition would have caused difficulties in latency measurements. Instead of the immediate repetitions, a magnetic click was sounded one second before the word was read in order to focus the child's attention to the oncoming stimulus.

Finally, the 10 second rate for stimulus presentation Anisfeld and Knapp, 1967, Underwood, 1965) was reduced to a 5-1/2 second interval. The single presentation of the stimulus instead of the repetition, and the verbal response required of the S rather than a written response made it possible to shorten the interstimulus interval. Furthermore, the shorter interval increased the likelihood that Ss would feel some pressure to respond quicker.

METHOD

Subjects

The sample was comprised of 80 children from the Ithaca School system. There were 40 third grade children that ranged in age from 8 years to 9 years 10 months, and 40 sixth graders between 10 years 11 months and 12 years 3 months. The mean age of the third graders was 8 years 6 months and the sixth graders, 11 years 5 months. An equal number of boys and girls were tested at each grade level. All subjects were native English speakers.

Word Selection

Each list was composed of (a) 12 target words; (b) 24 experimental words; (c) 24 control words; (d) 28 filler words. Each target word had two types of experimental words -- one phonetically related to it and one semantically related. The target and experimental words were selected from either the 4th grade norms of Jenkins and Palermo (1964) or from the third grade norms of Entwisle (1966). This provided a convenient access to a large sample of words with known association value. From this sample only those stimuli and their associates which appeared on the Watters and Courtis (1948) vocabulary list for 2nd graders were included in the study. The associates were then classified as synonyms or antonyms first, if they subjectively fit the category and secondly, if they were classified in Roget's

Thesaurus or Jenkins and Palermo (1965) norms as a synonym or antonym of the target word. The words were then classified as low or high associates. The mean association value for the high associates was .46 and the range was .21 to .38. The mean of the low associates was .07 and the range was .03 to .12. (See Table 1, p. 28 for specific values.)

The semantic experimental words were of four kinds: (a) three were synonyms of the target word but not highly associated with it as indicated by the Jenkins and Palermo (1964) or Entwisle (1966) word association norms (example, loud-noisy); (b) three were highly associated synonyms (example, house-home); (c) three, low associated antonyms (example, come-go); (d) and finally, three highly associated antonyms of the target word (example, hard-soft).

Words which were phonetically related to the target word were chosen arbitrarily if the author subjectively felt that the word endings rhymed. All but one of these phonetically related words appeared on the 2nd grade vocabulary list. The word "brief" was accidentally overlooked during word selection and a post-experiment check revealed that the 3rd grade Ss were unfamiliar with the meaning of the word.

Control words for the semantic and phonetic associates of the target words were selected on the following criteria: (a) no readily obvious relationship to any word in the list; (b) same part of speech as their experimental counterpart; (c) same number of syllables; and

Table 1. The frequency (F) of occurrence of the experimental and control words compiled from Thorndike-Lorge+ or Rinsland* word frequency counts and the association value (A.V.) between the target word and its semantic experimental counterpart as measured by Palermo and Jenkins or Entwisle word association norms

	Target word	A.V.	Semantic exper.	F	Semantic control	F	Rhyme exper.	F	Rhyme control	F
Hi Assoc.	List A									
	lamp	.64	light	M	bank	M	camp	380	flag	380
	mountain	.35	hill	M	street	M	fountain	153	hammer	157
	kitten	.39	cat	410	shell	410	mitten	8	thimble	8
	boy	.48	girl	M	fire	M	toy	221	priest	225
	slow	.30	fast	*246	next	*246	blow	* 77	hang	* 78
	up	.88	down	M	less	M	cup	700	chair	700
	List B									
	city	.44	town	M	life	M	kitty	* 45	zebra	* 45
	house	.21	home	M	ship	M	mouse	193	glove	194
	carpet	.53	rug	157	soup	166	market	700	army	700
	hard	.38	soft	M	cute	M	card	M	wind	M
	high	.51	law	M	first	M	my	M	her	M
	white	.39	black	M	round	M	right	M	best	M
Lo Assoc.	List A									
	thief	.04	crook	15	tag	21	brief	192	ripe	186
	loud	.05	noisy	92	crazy	98	cloud	* 26	boot	* 20
	wish	.10	want	M	look	M	fish	700	coat	700
	live	.10	die	M	bring	M	give	M	could	M
	cry	.04	laugh	M	learn	M	try	M	tell	M
	find	.03	lose	700	hurt	700	mind	M	king	M

Table 1. (Continued)

Target word	A.V.	Semantic exper.	F	Semantic control	F	Rhyme exper.	F	Rhyme control	F
List B									
Lo Assoc.									
gun	.06	rifle	132	bubble	129	run	M	cut	M
stove	.12	oven	30	shelf	31	drove	260	dig	260
cold	.05	freezing	* 3	cheerful	* 3	hold	M	use	M
go	.05	come	M	keep	M	row	450	feed	453
sweet	.11	sour	50	slump	59	feet	M	bird	M
take	.06	give	M	kill	M	make	M	help	M

+ Raw F represents the occurrence of the word per 4-1/2 million words and M indicates that the word occurred 1,000 times or more in the count of 120 juvenile books.

* Raw F represents the occurrence of the word per one hundred thousand words in Rinsland's norms.

(d) frequency of occurrence. Most of the frequency data were compiled from the Thorndike Lorge tables (1944) but in four cases frequency counts from Rinsland (1947) were employed. In all cases both the frequency data of the experimental and its control word were selected from the same source; (e) all control words had to appear on the Watters and Courtis (1948) vocabulary list.

The filler words were selected randomly from the Watters and Courtis vocabulary list and an attempt was made to choose words unrelated to other words in the list.

List Construction

There were two lists of 121 words each. Half of the Ss at each grade level were given list A and half list B. Each target word was presented twice in the list where the second token of the word followed the first by eight positions. Each of the experimental words was presented only once in the list as was its corresponding control word. The control words were placed one space away from their experimental counterparts. Eight times the control word followed its experimental word and 16 times it preceded the presentation of its experimental word. Filler words were distributed in the unused spaces. Ten filler words were presented only once, fifteen twice and three, three times. Filler words were repeated so that the Ss would actually hear the same word several times in the list. This was done to give the Ss a set for frequent repetition of words.

The pair of experimental words (the one phonetically related to the target word and the other semantically related) always followed the second presentation of the appropriate target word. One member of the experimental pair was separated from the target by nine spaces and the second member by fourteen spaces, leaving four positions between the presentation of the first member of the experimental pair and the second member. Half the time in the list the semantically related member of the pair was in the initial position with the phonetically related word in the latter place, and half the time the phonetically related word preceded the semantically related member. Two orders of the same list of words were constructed to control for the presentation of semantic and phonetic members. In order I, the semantically related experimental word preceded the phonetic member for the odd numbered target words in the lists. In order II, the semantic member was presented ahead of the phonetic member for the even numbered target words. (See Appendix, p.56-7, for schematic plan of lists.)

Each of the four types of semantic experimental words (high-associated synonyms, low-associated synonyms, high-associated antonyms, and low-associated antonyms) was distributed evenly across the list. That is, in each third of the list there was one of each of the four types of semantically related experimental words.

List A and list B contained different target, experimental and control words. But the filler words on both lists were the same. The two lists were constructed so that the same number of words from

the four categories of association appeared on both lists. Words were randomly assigned to lists but then reshuffled to insure that the words on one list were related only to their target word and not to any other word on the list. For example, "kitty," rhyme word of "city" had to be separated from "kitten," target word for "cat."

Procedure

The words were presented on a tape recorder at a rate of one word every 5-1/2 seconds. A second before the word was presented a warning "click" was given so that the S had 4-1/2 seconds to respond from the onset of the stimulus word. The total testing time was about 10 minutes.

Each S was tested individually and the session was recorded. The following instructions were given to each child:

> Soon you will be presented with a long list of words on this tape recorder. Some of the words you will hear only once in the list. But other words you will hear once and then you will hear them again later in the list. What you have to do is to listen carefully to each word and after each word say "new" if you are hearing the word for the first time in the list or say "old" if you heard the word earlier in the list. Try and say it as quickly as you can.

A pretest trial was given where the words were actually repeated a second time and no related words were presented. The purpose was to see if the child understood the instructions regarding responding. If an error was made on the series the child was corrected and a second pretrial was run. It was assumed that on this list errors would be due to a misunderstanding of the instructions and not errors due to intrusion.

After the session the tapes were played through a rectifier into a Brush Recorder. This process permitted the auditory input from the tape to activate a recording pen which then produced a visual presentation on graph paper. The pen was inactive when no sound was present on the recording, tracing a straight line on the recording paper which ended with a deflection corresponding to the onset of the S's response. By measuring the straight line between the onset of the stimulus word and the onset of the S's response, it was possible to determine a response latency measure to each stimulus word. Only entire word responses were included in the latency measure, all false starts were excluded as was the entire word which may have followed. Latencies were recorded in millimeters where: 2mm = 1 second.

Certain problems during testing interferedwith a strict adherence to the procedure. In a few instances the S did not hear the stimulus word and asked that the experimenter repat it. In these cases the word was repeated and the S's response was included in the analysis of errors but not in the latency analysis. Four Ss failed to respond to one or two list words; they were given the benefit of the doubt and the blank responses were not scored a errors.

Two Ss were excluded from the sample before an examination of their distribution of errors to experimental and corrol words. One because his error rate was so much in excess of the average error rate for Ss. It was felt that his responses might be unreliable. The other S became confused with the task and on hearing an experimental word

verbalized the fact that he had the synonym previously and asked whether the present word should then be judged as "new" or "old." The tape was stopped and the instructions given again. He completed the task with no more difficulty -- responding correctly to most of the words and making the average amount of intrusion errors.

When the experiment was completed, the lists were played for a different group of Ss in order to determine 1) any misperceptions of the list words and 2) whether 3rd grade Ss were familiar with the test words. Twelve 3rd graders and 12 sixth graders listened to either list A or list B. The third graders had to repeat each word and use it in a sentence or define it. The sixth graders wrote down the words they heard.

Five of the six third grade Ss presented with the word "brief" were unfamiliar with it. But all other words could be used by the third graders. The following misperceptions were made by both the 3rd and 6th graders:

(1) List A, target and experimental words - camp (1S - tamp), live (4 - leave), die (2 - dye), thief (1 - beef, 1 - feef, 1 - feet), brief (1 - brieth), laugh (1 - lac); control words - chair (1 - share).

(2) List B, target and experimental words - cold (3Ss - called), high (5 - hi), my (2 - mine), sweet (1 - swept); control words - first (1 - burst), life (1 - like), her (1 - purr), plump (1 - clump), glove (1 - clove), wind (1 - winged), bubble (1 - ?).

RESULTS

Findings on the Differences Between Experimental and Control Words

In evaluating the recognition errors made, it should be noted that the amount of errors of any type was minimal. The third graders made errors to 9% of the words on the list, while the 6th graders made errors to 8% of the words. The total number of errors to the first presentation of the target words was 36 for the third graders and 27 for the sixth graders; the totals for the second presentation of the target words was 43 and 21, respectively. These sums can be compared with the errors to experimental and control words presented in Table 2. For both grades, the total number of errors to the experimental words was larger than the number of errors to control words.

Age differences were found when errors to the different types of experimental words were compared with the errors to the respective control words. (See Table 3) The sixth grade made significantly more errors to all the five categories of experimental words than to the respective controls. However, for the third grade only the errors to the Rhyme words and to the Hi Assoc. Synonyms were significantly greater than errors to the corresponding control words. For the other 3 word categories there was no significant difference between errors to experimental and control words in the 3rd grade group of Ss.

Differences in the reaction time of responding correctly to the experimental and control words were compared. To reduce the large

Table 2. Total number of errors made by the third and sixth graders to each word

	Semantic Exper.	Errors 3rd	Errors 6th	Semantic Control	Errors 3rd	Errors 6th	Rhyme Exper.	Errors 3rd	Errors 6th	Rhyme Control	Errors 3rd	Errors 6th
Hi Ass. Syn.	List A											
	light	3	1	bank	1	1	camp	3	0	flag	0	0
	hill	1	2	street	1	1	fountain	1	3	hammer	3	4
	cat	3	3	shell	1	1	mitten	7	3	thimble	1	1
	List B											
	town	1	1	life	0	0	kitty	0	1	zebra	0	0
	home	5	7	ship	1	2	mouse	4	2	glove	1	3
	rug	8	6	soup	4	1	market	3	4	army	0	1
Lo Ass. Syn.	List A											
	crook	4	3	tag	1	0	brief	1	2	ripe	0	0
	want	3	2	look	6	3	fish	4	1	coat	2	0
	noisy	6	7	crazy	2	0	cloud	4	2	boot	0	1
	List B											
	rifle	1	2	bubble	2	0	run	3	5	cut	0	0
	freez-ing	0	0	cheerful	3	3	hold	5	5	use	4	2
	oven	5	8	shelf	3	4	drove	1	2	dig	0	1
Hi Ass. Opp.	List A											
	girl	0	1	fire	0	0	toy	2	2	priest	2	1
	fast	7	6	next	2	0	blow	4	0	hang	3	4
	down	4	7	less	3	1	cup	4	4	chair	4	4
	List B											
	soft	0	2	cute	2	2	card	3	3	wind	2	0
	low	1	3	first	0	3	my	1	3	her	2	2
	black	2	1	round	4	2	right	2	2	best	0	4
Lo Ass. Opp.	List A											
	die	0	1	bring	2	0	give	2	4	could	1	0
	laugh	4	9	learn	2	2	try	9	10	tell	3	0
	lose	6	5	hurt	6	3	mind	3	2	king	2	0
	List B											
	come	1	4	keep	1	1	row	3	1	feed	0	1
	sour	1	2	plump	0	1	feet	8	3	bird	1	2
	give	6	9	kill	0	2	make	6	6	help	6	4
Grand Total		72	92		47	33		83	70		37	35
Mean		1.80	2.30		1.18	.82		2.08	1.80		.92	.88

Table 3. Mean number of errors to the different types of experimental and control words by age level

	Rhyme (24 words)	Semantic Hi Assoc. Synon. (6 words)	Semantic Hi Assoc. Anton. (6 words)	Semantic Lo Assoc. Synon. (6 words)	Semantic Lo Assoc. Anton. (6 words)
Grade 3					
Exper.	2.08 t=3.71*	.52 t=3.35*	.35 t < 1	.48 t < 1	.45 t < 1
Control	.92	.20	.28	.42	.28
Grade 6					
Exper.	1.80 t=2.66**	.50 t=3.02*	.50 t=2.08**	.55 t=2.51**	.75 t=3.41*
Control	.88	.15	.20	.25	.22

* $p < .01$
** $p < .05$

variability in an individual's latency data the practice of using the median latency as the score rather than the S's mean was employed. The mean of the (median) reaction time scores to experimental words was 1.56 seconds for the 6th graders and 1.82 seconds for the third graders compared to 1.46 and 1.70 seconds respectively to the control words. An analysis of variance showed that the reaction time was significantly longer to the experimental words than to the control words for both third and sixth graders (F=17.99, df 3/192, $p < .001$). (See Appendix, Table 9 for a summary of the analysis of variance.)

A check was made to insure against the possibility that the experimental words might be longer than the controls and thus responsible for the longer reaction time obtained to the experimental words. The mean length of experimental words measured on the brush recorder tape was .64 seconds and the mean for controls was .63 seconds.

Furthermore, a scatter plot revealed no relationship between the length of the experimental word and the latency of the response. (See Appendix, Figure 3.) Thus, differential word length could not be responsible for the differences in latency to experimental and to control words.

It should be noted that this analysis of reaction time indicated significant sex differences in the speed of responding. As can be seen from Table 4 it took longer for boys to respond than girls, at both age levels. These differences are significant ($p < .05$). (See Analysis of Variance in Appendix, Table 7.) No apparent explanation could be found to support this difference since the research in the area of verbal fluency offer contradictory findings with respect to sex differences.

Table 4. Mean reaction time score for boys and girls

	Boys	Girls
3rd grade	1.80	1.72
6th grade	1.59	1.43
Total	1.69	1.59

Findings on the Differences Between the Various Types of Experimental Words

Developmental differences were found between the number of errors made to the different experimental word types. For this analysis, error difference scores were used. An error difference score is defined as the number of errors to the experimental words less the errors to

the respective controls. As shown in Figure 1, there was an interaction between grade level and number of errors to the different word types. An analysis of variance revealed that this interaction was significant at $p < .025$. (See Appendix, Table 10 for a summary of the analysis.)

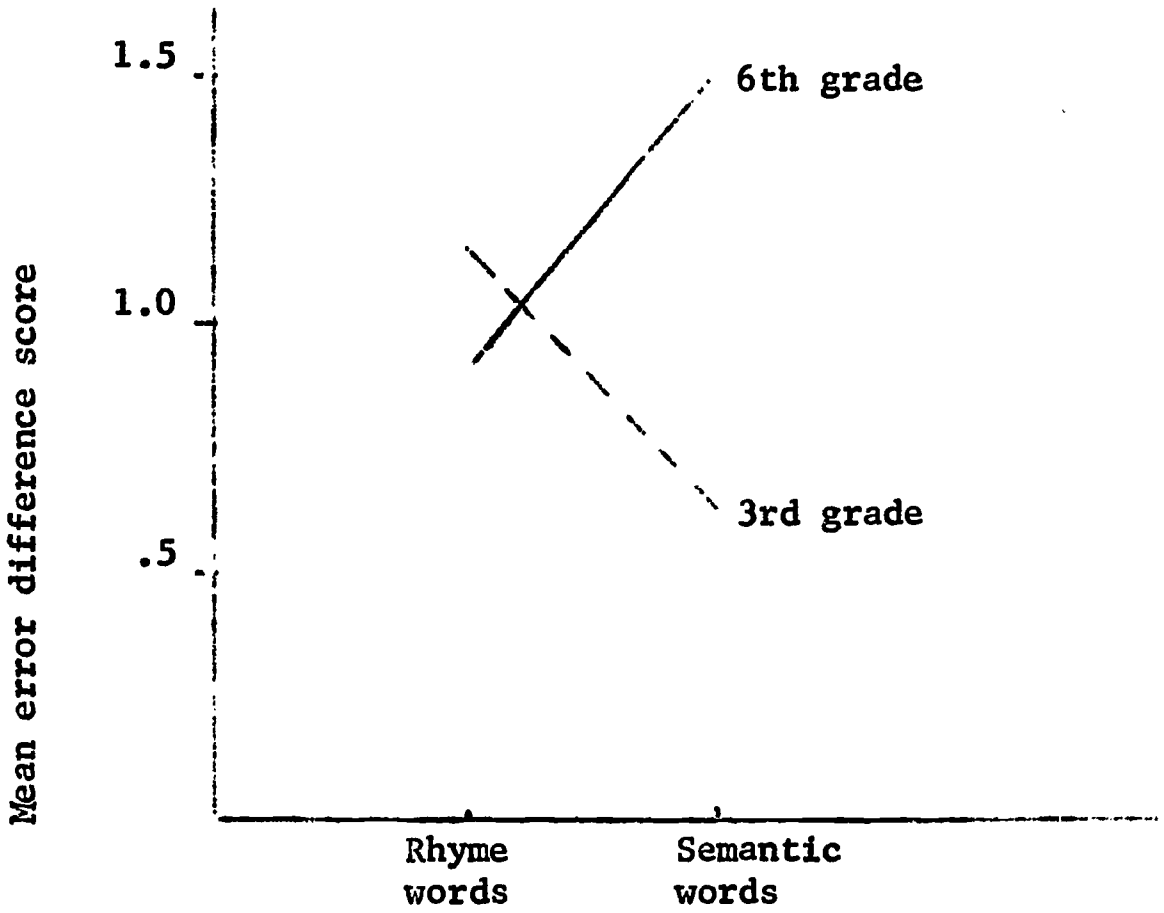

Figure 1. Differential amount of errors to Rhyme and Semantic words by grade level.

A Scheffe test was used to determine which comparisons were responsible for the interaction. A difference of .22 between pairs was needed for an F ratio to be significant at the .05 level. All comparisons were within the range of .24- .86 and thus significant. (See Appendix, Table 11, for the exact means and differences between pairs.) Thus, significantly more errors were made by third graders to the rhyme words than to the semantic words, while sixth graders made significantly more errors to semantic words than to rhyme words. Furthermore, third graders made significantly more errors to the rhyme words than did the sixth graders, while significantly more errors were made to the semantic words by the sixth graders than by the third graders.

An evaluation was made of the differential effects of synonyms and antonyms and of high and low associates in the production of errors. The difference score to synonyms was .38 for 3rd graders and .65 for 6th graders. While to antonyms the scores were .25 and .83 respectively. The results of an analysis of variance did not reveal any significant differences in the errors to these word types, either within or between grades. (See Appendix, Table 12.) But there seemed to be a tendency for 6th graders to make more errors to the antonyms than did 3rd graders while in errors to synonyms the developmental difference was less striking.

Similarily, no significant differences within or across grades were found in the amount of errors accounted for by Hi and Lo Assoc. (See Table 5.) However, the error difference score to Low Assoc. tended to be higher than to Hi Assoc. for the sixth graders, while the

third grade error difference score was higher to the Hi Assoc. than to the Low Assoc.

Table 5. Mean error difference scores to Hi and Low Assoc. by grade

	Hi Assoc.	Low Assoc.
3rd grade	.39	.23
6th grade	.65	.83
3rd and 6th grade	.52	.53

The analysis of variance of differences between Hi and Low Assoc. Synonyms and Antonyms indicated a significant interaction effect between lists and the number of errors to the different semantic word types ($F = 6.47$, df 3/216, $p < .01$). (See Appendix, Table 13, for a summary of the analysis.) As shown in Table 1, the word types contributing to most of the errors on one list accounted for the least on the second list. (Also See Table 6).

Table 6. Mean error difference scores to the Semantic experimental words by list

	Hi Assoc.		Low Assoc.	
	Synonyms	Antonyms	Synonyms	Antonyms
List A	.17	.47	.32	.25
List B	.50	- .10	.02	.45

Considering that each subgroup contained only three words, it is not surprising that pecularities of the individual words should play a role. For instance, the lessened effectiveness of the Hi Assoc. Antonyms to produce errors on list B, may have been a consequence of misinterpretations of the target words. Ss asked to listen to the tape, after the sample was tested, intepreted the target word 'high' as 'hi' on list B. This could likely reduce the number of false recognition errors one would expect to the experimental word 'low' on this list.

The distribution of latency scores was evaluated for the different experimental words and grade levels. The mean reaction time was 1.69 seconds to the phonetic words and 1.69 seconds to the semantic words. The latency scores differed only between ages across word types, with a significantly shorter response time from third graders to all words. The third grade mean reaction time score was 1.82 seconds to both rhyme and semantic words, while the sixth grade mean was 1.56 seconds. These age differences were significant at the .001 level (See Appendix, Table 9) and may reflect the expected verbal facility of 6th graders over third graders.

Significant differences were found in reaction time between the words on the different lists with a mean reaction time on list A of 1.65 seconds compared to a mean latency of 1.62 seconds on list B. An analysis of variance indicated that the interaction was significant at .001 level (See Appendix, Table 9). A Scheffe test required the

pair difference to be greater than .12 seconds to be significant. The range of obtained differences was .03-.15. Only the difference between the semantic control words on the two lists was responsible for the interaction effect. Since there were no significant differences in the reaction time to the experimental words on the two lists, one may assume that certain unforseen characteristics of the control words were responsible for the difference.

Differences in Reaction Time for Correct and Incorrect Responses

An analysis of variance was performed on the latency differences between correctly and falsely recognizing an experimental word. The analysis was limited to the experimental words because of the interest in studying words related to other words and not in the errors to unrelated controls. Only Ss that had made errors to both rhyme and semantic words and therefore, had a median latency score for incorrect responses, were included in the analysis. Additional Ss were randomly excluded to maintain a balanced design thus limiting the N to a total of 40 Ss, 20 per grade. The reaction time for correct responses was 1.71 seconds and 1.93 seconds for incorrect responses. This was significant at the .001 level (See Appendix, Table 14, for the analysis). Moreover, it took significantly longer to make an error than to respond correctly, i.e., it takes longer to respond "old" than "new" to an experimental word. This pattern holds for both grades (See Figure 2). It should be noted that in some cases the median latency score for the incorrect responses was based on only one response which tends to weaken the results based on these scores.

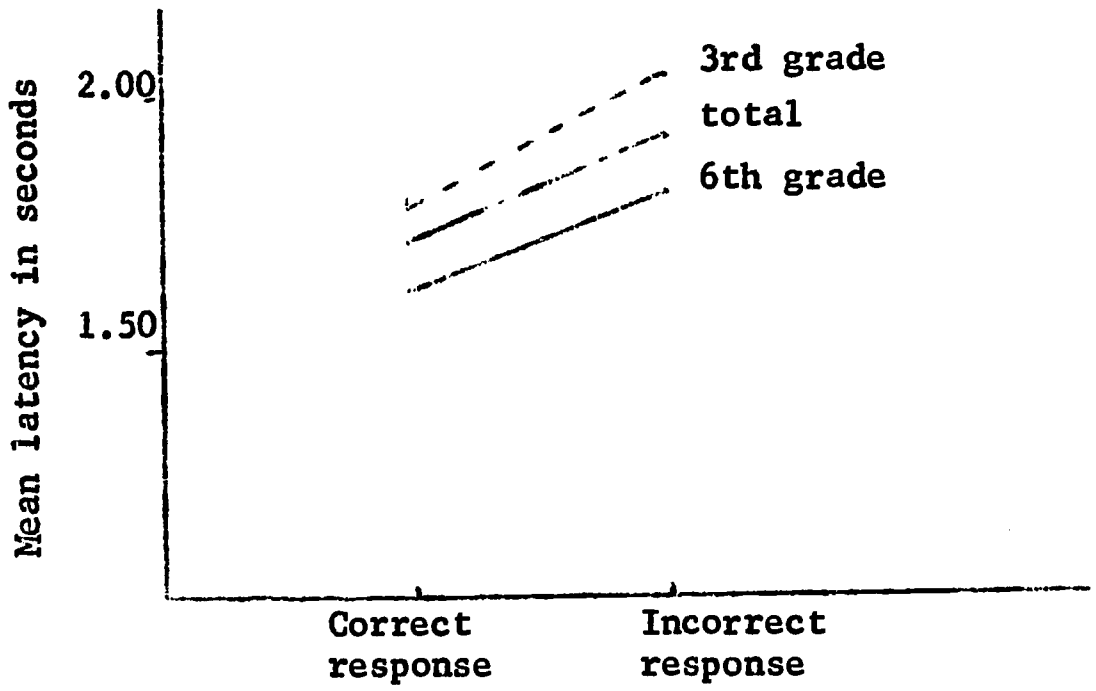

Figure 2. Mean latency responses to experimental words by correctness of response.

Also it took significantly longer to make an error on list A, mean 2.02 seconds, than on list B, mean 1.45 seconds. This is probably attributable to certain unexplainable list differences.

Summary

These data show that Ss were much more likely to recognize a word correctly as new or old than to make an error and that it took longer to make an error than to respond correctly. In evaluating the errors, it was seen that Ss made significantly more errors to experimental words than to controls and took longer to respond to the experimentals than to the controls. Developmental differences existed in the distribution of errors. Sixth graders made signifi-

cantly more errors to all categories of the experimental words than to the controls, but for third graders these experimental minus control word differences were significant only for the Rhyme and Hi Assoc. Synonyms.

Furthermore, these results indicate that the third graders made significantly more errors to rhyme words than did sixth graders, while the sixth graders made more errors to semantic words than did the third graders. Furthermore, third graders made more errors to rhyme words than to semantic words and the sixth graders made more errors to the semantic words than to the rhyme words. No significant differences were found in the amount of errors to synonyms or antonyms or to Hi and Low Assoc. Certain developmental trends were, however, suggested. The third graders make more errors to synonyms than to antonyms and more errors to Hi Assoc. than to Low Assoc. It is the antonyms and Low Assoc. which account for most of the errors for the sixth graders.

DISCUSSION

The occurrence of significantly more errors to experimental words than to control words supports the results of other studies of false recognition memory (Anisfeld and Knapp, 1967; Underwood, 1965). Any interpretation of these findings must account both for the occurrence of false recognition errors as well as the S's high probability of correctly recognizing a word on the list. The false recognition errors to experimental words suggest the existence of process whereby words are grouped into clusters of related words. Since the experimental words are semantically or phonetically related to other words in the list, it is likely that at the time of its presentation the experimental word might be linked with a word heard previously in the list. The control words, on the other hand were chosen specifically because of no apparent relationship with any other word in the list. It is therefore, unlikely that when the control word is presented that it will be grouped with another word from the earlier part of the list.

The results of this study as well as previous false recognition studies suggest that Ss are much more likely to recognize a word correctly than to make an error. Such a finding suggests that even though there seems to be a process of grouping related words with each other, each word maintains its separate identity. For example, 'house' and 'home' are related as synonyms but each has a different

affectional and referential meaning which helps to maintain the word's distinct identity in the group of related words.

The findings reported suggest that the psychological relationships among words differ as a function of age. This study like previous developmental studies has demonstrated an age gradient with respect to the semantic and phonetic relations between words. With increasing age there appears to be an increase in the effectiveness of the semantic relations as a basis for grouping words. Also, the results are consistent with Riess and Luria that with the older Ss there is a reduced effect of the phonetic relations. It appears that the younger child groups words with other words, for the most part, on the basis of physical similarity of words, namely common sound features. The younger child may be familiar with the synonyms and antonym of a particular word but the labels have not been fully internalized in the meaning system. The older child groups words according to semantic relations of similarity based on meaning systems acquired through experience with language.

There were no significant differences in the amount of errors made between synonyms and antonyms or between the high and low associates either within or across grades. These findings are discrepant with the developmental gradient reported by Riess where the younger Ss generalized significantly more to antonyms than to synonyms while the older Ss generalized more to the synonyms than to antonyms.

Also, these data are at variance with the emphasis placed on association value in the clustering studies and in Underwood's study of false recognition memory. In the present study low associates contributed to the production of false recognition errors not significantly less than Hi Associates. It should be noted that these findings are in line with Luria's description of semantic generalization in the retarded child. Also the results are congruent with Anisfeld and Knapp's (1967) conclusion that logical relations per se and not association values alone are important for the organization of words.

A developmental trend, though nonsignificant was shown where younger Ss responded more to the synonyms than to the antonyms and more to the Hi Assoc. than to the Low Assoc. Furthermore, the young child made significantly more errors only to the semantic experimental words representing the Hi Assoc. Synonym category than to the respective controls. Such findings support Luria's description of the severely retarded child who will emit the orienting response only to a semantic word having a very close 'connection' with the signal word. These observations may be a consequence of the young child's overall level of cognitive development at this age. Due to his still limited experience with words, he may have an incomplete understanding of the meaning of a word. The young child may be able to relate a word to only a small range of semantically similar words.

It is often assumed that the organizing of words with other words occurs in the storage phase of memory and that related words

are stored with one another thereby reducing the number of storage units without the consequent loss of information. However, it is conceivable that the process occurs rather during recall where there may be a restructuring of words so related words are associated together. The functioning of the associative process during recall may explain the difference in reaction time that were reported.

The latency to respond correctly to an experimental word is longer than to a control word. When the S hears the experimental word he may at that time associate it with a word heard earlier in the list. Yet, since the experimental word was actually not presented before, and since each word maintains its own identity even though the process of association occurs, the S is likely to recognize the word as 'new'. It is just the momentary association with an earlier word which increases the S's uncertainty during recall and leads to a longer reaction time. At the presentation of the control word the S is not likely to be confronted with the possibility that he might have heard the word before in the list therefore, the S can respond faster to the control word than to the experimental word.

The longer reaction time to experimental words for incorrect responses than for correct responses suggests several explanations. First, it just may take a S longer to give a negative response, that is, to say "no" or in this case "old" than to say "yes", or "new" as required in this task. Such a possibility is suggested by those studies where the S tends to agree more readily than to disagree

especially when there is some ambiguity in the stimulus (Peabody, 1961). From the data it is not possible to evaluate this position but it does suggest the need for further investigation with an appropriate experimental design. Second, since a large majority of the words on the list appeared only once, the S may have become accustomed to giving the "new" response and when contemplating saying "old" was slowed down. Third, it is possible that the longer latency for incorrect responses than for correct responses is due to the relations built into the list. When the S forms a weak link between the experimental word and an earlier word in the list such that the experimental word maintains its separate identity, then he responds to it as a 'new' word and the reaction time is fast. However, when the relationship with an earlier word becomes stronger and the word loses its distinctiveness in the cluster of related words, the S's uncertainty about the particular word increases and he gives an incorrect response and is slowed down.

SUMMARY

This study was conducted to explore developmentally the features of words used in the process of organizing related words together. The experimental design called for a memory task in which inferences about the way Ss organized related words was based on an evaluation of recognition errors where Ss indicated that a word associated with an earlier word had been presented before.

Forty 3rd and forty 6th graders were presented aurally with a list of 121 words. The list was composed of experimental words which were phonetically or semantically related to a word presented earlier in the list (target word). One position removed from each experimental word was a control word matched in syllablic length, frequence of occurrence and part of speech. The control words had no apparent relationship to any other word in the list and thus provided a comparison of errors to words neither semantically nor phonetically related to earlier presented list words.

Each target word was followed by both a phonetically and semantically related word, each presented only one time in the list. The phonetically related word rhymed with its previously presented target word while the semantic relationship was defined by the dimensions of association and synonymity or antonymity. The Hi Associates

were chosen from Palermo and Jenkins and Entwisle's norms and the range of association value was .21-.88. The Lo Associates ranged in association value from .03-.12. Half of the Associates in each group were synonyms and half antonyms.

The S's task was to indicate at the presentation of each word whether the word had been read before ("old") in the list or not ("new"). In particular, this study was concerned with the intrusion or false recognition errors where a word was recognized as having been presented before. The errors score was supplemented by a measure of the S's reaction time for responding.

The results show that Ss were much more likely to recognize a word correctly than to make an error. An evaluation of errors indicated that Ss made significantly more recognition errors to experimental words than to controls and an analysis of latency results showed that it took longer to respond to experimental words than to controls. Furthermore, it was found that it took significantly longer to make an error than to respond correctly to the experimental words. Developmental differences were revealed in the distribution of errors. Third graders made significantly more errors to rhyme words than to semantic related words while the 6th graders made more errors to the semantic than to the rhyme words. No significant differences were found in the amount of errors to synonyms and antonyms or to Hi and Lo Associates.

These findings were interpreted as suggestive of a process whereby related words are grouped together so that at the presentation of an experimental word, the S might link it with the word to which it is related and which was heard previously in the list. It is this process which explains the larger number of errors and the longer reaction time to experimental words than to controls. The developmental differences suggest that different relationships between words are used for the grouping of words at the different age levels. Besides this grouping process the Ss high probability of correctly recognizing a word may indicate that words maintain their separate identity and can thus be correctly identified.

REFERENCES

Anisfeld, M., & Knapp, M. Association, synonymity and directionality in false recognition. (In press)

Bousfield, W. The occurrence of clustering in the recall of randomly arranged associates. Journal of General Psychology, 1953, 49, 299-240.

Bousfield, W., Cohen, B., & Whitmarsh, G. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. Psychological Reports, 1958, 4, 39-44.

Cofer, C.N. A study of clustering in free recall based on synonyms. Journal of General Psychology, 1959, 60, 3-10.

Entwisle, D. Word associations of young children. Baltimore: Johns Hopkins Press, 1966.

Feather, B.W. Semantic generalization of classically conditioned responses: a review. Psychological Bulletin, 1965, 63, 425-441.

Hartman, T.F. Semantic transfer of the differential conditioned eyelid response from words to objects. Journal of Experimental Psychology, 1963, 65, 194-200.

Jenkins, J.J., Mink, W.D., & Russell, W.A. Associative clustering as a function of verbal association strength. Psychological Reports, 1958, 4, 127-136.

Jenkins, J. & Palermo, D. Word association norms: grade school through college. Minneapolis: University of Minnesota Press, 1964.

Lacey, J.I., & Smith, R.L. Conditioning and generalization of unconscious anxiety. Science, 1954, 120, 1045-1052.

58.

Luria, A.R. The mentally retarded child. New York: Pergamon Press, MacMillan Co., 1963.

Olver, R. A development study of cognitive equivalence. Unpublished doctoral dissertation, Radcliffe College, 1961.

Peabody, D. Attitude content and acquiescent set. Journal of Abnormal and Social Psychology, 1961, 63, 1-11.

Razran, G.A. A quantitative study of meaning by a conditioned salivary technique. Science, 1939. 90, 89-90.

Rice, U. A developmental study of semantic and phonetic generalization in paired-associate learning. Unpublished doctoral dissertation, Syracuse University, 1963.

Rice, U.M., & DiVesta, F.J. A developmental study of semantic and phonetic generalization in paired associate learning. Child Development, 1965, 36, 721-730.

Riess, B.F. Semantic conditioning involving the galvanic skin reflex. Journal of Experimental Psychology, 1940, 26, 238-240.

Riess, B.F. Genetic changes in semantic conditioning. Journal of Experimental Psychology, 1946, 36, 143-152.

Rinsland, H. A basic vocabulary of elementary school children. New York: MacMillan Co., 1947.

Shepard, R.N., & Teghtoosian, M. Retention of information under conditions approaching a steady state. Journal of Experimental Psychology, 1961, 62, 302-309.

Sigel, I. The attainment of concepts. In M. Hoffman & L. Hoffman (Eds.), Review of child development research. Vol. 1. New York: Russell Sage Foundation, 1964. Pp. 209-249.

Thorndike, E. & Lorge, I. The teacher's word book of 30,000 words. New York: Teacher's College, Columbia University, 1944.

Underwood, B.J. False recognition produced by implicit verbal responses. Journal of Experimental Psychology, 1965, 70, 122-129.

Watters, G. & Courtis, S.A. The picture dictionary for children. New York: Grosset and Dunlap, 1948.

Woodworth, R. Experimental psychology. New York: H. Holt & Co., 1958.

APPENDIX

Table 7. Schematic presentation of Order I

1.	Filler words
2.	
3.	
4.	
5.	
6.	Target word a^1
7.	Target word b^1
8.	Filler words
9.	
10.	
11.	
12.	
13.	Target word c^1
14.	Target word a^2
15.	Target word b^2
16.	Filler words
17.	
18.	
19.	
20.	
21.	Target word c^2
22.	Control Sem. word a
23.	Control Rhyme word b
24.	Semantic Exper. word a
25.	Rhyme Exper. word b
26.	Target word d^1
27.	Control Rhyme word a
28.	Control Sem. word b
29.	Rhyme Exper. word a
30.	Semantic Exper. word b
31.	Semantic Exper. word c
32.	Filler word
33.	Control Sem. word c
34.	Target word d^2
35.	Filler word
36.	Rhyme Exper. word c
37.	Filler
38.	Control Rhyme word c
39.	Target word e^1
40.	Target word f^1
41.	Filler words
42.	
43.	
44.	Rhyme Exper. word d
45.	Filler word
46.	Control Rhyme word d
47.	Target word e^2
48.	Target word f^2
49.	Semantic Exper. word d
50.	Filler
51.	Control Sem. word d
52.	Filler words
53.	
54.	
55.	Control Sem. word e
56.	Control Rhyme word f
57.	Semantic Exper. word e
58.	Rhyme Exper. word f
59.	Filler word
60.	Control Rhyme word e
61.	Control Sem. word f
62.	Rhyme Exper. word e
63.	Semantic Exper. word f
64-121.	Repeats pattern 6-63, with target words g-L

Table 8. Schematic presentation of Order II

1.	Filler words
2.	
3.	
4.	
5.	
6.	Target word a^1
7.	Target word b^1
8.	Filler words
9.	
10.	
11.	
12.	
13.	Target word c^1
14.	Target word a^2
15.	Target word b^2
16.	Filler words
17.	
18.	
19.	
20.	
21.	Target word c^2
22.	Control Rhyme word a
23.	Control Semantic word b
24.	Rhyme Exper. word a
25.	Semantic Exper. word b
26.	Target word d^1
27.	Control Sem. word a
28.	Control Rhyme word b
29.	Semantic Exper. word a
30.	Rhyme Exper. word b
31.	Rhyme Exper. word c
32.	Filler word
33.	Control Rhyme word c
34.	Target word d^2
35.	Filler word
36.	Semantic Exper. word c
37.	Filler word
38.	Control Semantic word c
39.	Target word e^1
40.	Target word f^1
41.	Filler words
42.	
43.	
44.	Semantic Exper. word d
45.	Filler word
46.	Control Sem. d
47.	Target word e^2
48.	Target word f^2
49.	Rhyme Exper. word d
50.	Filler word
51.	Control Rhyme word d
52.	Filler
53.	
54.	
55.	Control Rhyme word e
56.	Control Sem. word f
57.	Rhyme Exper. word e
58.	Semantic Exper. word f
59.	Filler word
60.	Control Sem. word e
61.	Control Rhyme word f
62.	Semantic Exper. word e
63.	Rhyme Exper. word f
	etc.

64.-121. Repeats pattern 6-63, with target words g-L.

Table 9. Six way analysis of variance of the individual's latency response to experimental and control words for correct responses

Source	Sum of squares	df	Mean square	F	P
Grade	3,165.76	1	3,165.76	23.23	.001
List	68.91	1	68.91	< 1	--
Sex	642.70	1	642.70	4.72	.05
Order	70.79	2	35.39	< 1	--
Grade x list	313.04	1	313.04	2.30	n.s.
Grade x sex	78.51	1	78.51	< 1	--
Grade x order	371.08	2	185.54	1.36	n.s.
List x sex	3.51	1	3.51	< 1	
Sex x order	83.14	2	41.57	< 1	--
Grade x list x sex	12.60	1	12.60	< 1	--
Grade x sex x order	31.51	2	15.75	< 1	--
Subject between error	8,720.67	64	136.26	--	--
Word type	663.88	3	221.29	17.99	.001
Word x grade	5.63	3	1.88	< 1	--
Word x list	241.79	3	80.60	6.55	.001
Word x sex	12.75	3	4.25	< 1	--
Word x order	111.59	6	18.60	1.51	n.s.
Word x grade x list	21.28	3	7.09	< 1	--
Word x grade x sex	47.60	3	15.86	1.29	n.s.
Word x list x sex	34.11	3	11.37	< 1	--
Word x grade x order	74.27	6	12.38	1.01	n.s.
Word x sex x order	30.22	6	5.04	< 1	--
Word x grade x list x sex	29.38	3	9.76	< 1	--
Word x grade x sex x order	65.72	6	10.95	< 1	--
Subject x word (error within)	2,361.32	192	12.30	--	--

Table 10. Six way analysis of variance of a comparison of the errors to rhyme and semantic experimental words

Source	Sum of squares	df	Mean square	F	P
Grade	3.31	1	3.31	< 1	--
List	1.41	1	1.41	< 1	--
Sex	.06	1	.06	< 1	--
Order	3.46	2	1.73	< 1	--
Grade x list	.51	1	.51	< 1	--
Grade x sex	3.31	1	3.31	< 1	--
Grade x order	13.41	2	6.71	1.59	n.s.
List x sex	.06	1	.06	< 1	--
Sex x order	3.81	2	1.91	< 1	--
Grade x list x sex	1.06	1	1.06	< 1	--
Grade x sex x order	.76	2	.38	< 1	--
Error between	246.20	64	3.85	--	--
Word type	.06	1	.06	< 1	--
Word x grade	12.66	1	12.66	5.35	.025
Word x list	1.06	1	1.06	< 1	--
Word x sex	1.06	1	1.06	< 1	--
Word x order	.11	2	.06	< 1	--
Word x grade x list	.31	1	.31	< 1	--
Word x grade x sex	1.06	1	1.06	< 1	--
Word x list x sex	3.91	1	3.91	1.65	n.s.
Word x grade x order	3.06	2	1.53	< 1	--
Word x sex x order	9.56	2	4.78	2.02	n.s.
Word x grade x list x sex	2.76	1	2.76	1.16	n.s.
Word x grade x order x sex	.51	2	.26	< 1	--
Error within	151.40	64	2.36	--	--

Table 11. Results of the Scheffé Test to evaluate the significant interaction between grade and number of errors to the different word types

Source	Mean error difference scores	Obtained pair difference
1. Rhyme words		
3rd grade	1.16	.24*
6th grade	.92	
2. Semantic words		
3rd grade	.62	.86*
6th grade	1.48	
3. Third grade		
Rhyme words	1.16	.54*
Semantic words	.62	
4. Sixth grade		
Rhyme words	.92	.56*
Semantic words	1.48	

* P < .05

Table 12. Five way analysis of variance of a comparison of the error to synonyms and antonyms

Source	Sum of squares	df	Mean square	F	P
Grade	7.22	1	7.22	4.84	.05
List	1.22	1	1.22	< 1	--
Order	.72	2	.36	< 1	--
Grade x list	.40	1	.40	< 1	--
Grade x order	5.82	2	2.90	1.88	n.s.
Error between	107.50	72	1.49	--	--
Word type	.02	1	.02	< 1	--
Word x grade	.90	1	.90	1.02	n.s.
Word x list	1.60	1	1.60	1.82	n.s.
Word x order	3.42	2	1.71	1.94	n.s.
Word x grade x list	1.22	1	1.22	1.39	n.s.
Word x grade x order	2.52	2	1.26	1.43	n.s.
Error within	6.38	72	.88	--	--

Table 13. Five way analysis of variance of the errors to high and low associated synonyms and antonyms

Source	Sum of squares	df	Mean square	F	P
Grade	3.61	1	3.61	4.84	.05
List	.61	1	.61	< 1	--
Order	.36	2	.18	< 1	--
Grade x list	.20	1	.20	< 1	--
Grade x order	2.91	2	1.46	1.94	n.s.
Error between	53.75	72	.75	--	--
Word type	2.12	3	.71	1.28	n.s.
Word x grade	1.11	3	.37	< 1	--
Word x list	10.71	3	3.57	6.47	.01
Word x order	2.54	6	..42	< 1	--
Word x grade x list	.68	3	.22	< 1	--
Word x grade x order	6.09	6	1.01	1.84	n.s.
Error within	119.25	216	.55	--	--

Table 14. Six way analysis of variance of the difference in latency response to experimental words when the response is correct and when the response is incorrect

Source	Sum of squares	df	Mean square	F	P
Grade	719.53	1	719.53	6.44	.05
List	81.65	1	81.65	<1	--
Sex	34.50	1	34.50	<1	--
Grade x list	217.86	1	217.86	1.95	n.s.
Grade x sex	37.34	1	37.34	<1	--
List x sex	18.02	1	18.02	<1	--
Grade x list x sex	.28	1	.28	<1	--
Error between	3,573.21	32	111.66	--	--
Word type (semantic or rhyme)	179.14	1	179.14	2.26	n.s.
Word x grade	67.21	1	67.21	<1	--
Word x list	69.30	1	69.30	<1	--
Word x sex	.06	1	.06	<1	--
Word x grade x list	116.11	1	116.11	1.46	n.s.
Word x grade x sex	.66	1	.66	<1	--
Word x list x sex	20.52	1	20.52	<1	--
Word x grade x list x sex	28.64	1	28.64	<1	--
Error within (subject x word interaction)	2,535.66	32	79.24	--	--
Correctness of response	1,241.55	1	1,241.55	31.61	.001
Correctness x grade	1.74	1	1.74	<1	--
Correctness x list	343.10	1	343.10	8.74	.01
Correctness x sex	3.11	1	3.11	<1	--
Correctness x grade x list	9.65	1	9.65	<1	--
Correctness x grade x sex	44.84	1	44.84	1.14	n.s.
Correctness x list x sex	19.39	1	19.39	<1	--
Correctness x grade s list x sex	4.00	1	4.00	<1	--
Error within (subject x correctness interaction)	1,256.88	32	39.29	<1	--
Word x correctness	74.66	1	74.66	1.57	n.s.
Word x correctness x grade	2.94	1	2.94	<1	--
Word x correctness x list	35.44	1	35.44	<1	--
Word x correctness x sex	42.74	1	42.74	<1	--
Word x correctness x grade x list	19.81	1	19.81	<1	--
Word x correctness x grade x sex	81.65	1	81.65	1.72	n.s.
Word x correctness x list x sex	37.34	1	37.34	<1	--
Word x correctness x grade x list x sex	10.87	1	10.87	<1	--
Error within (subject x word x correctness interaction)	1,518.45	32	47.45		

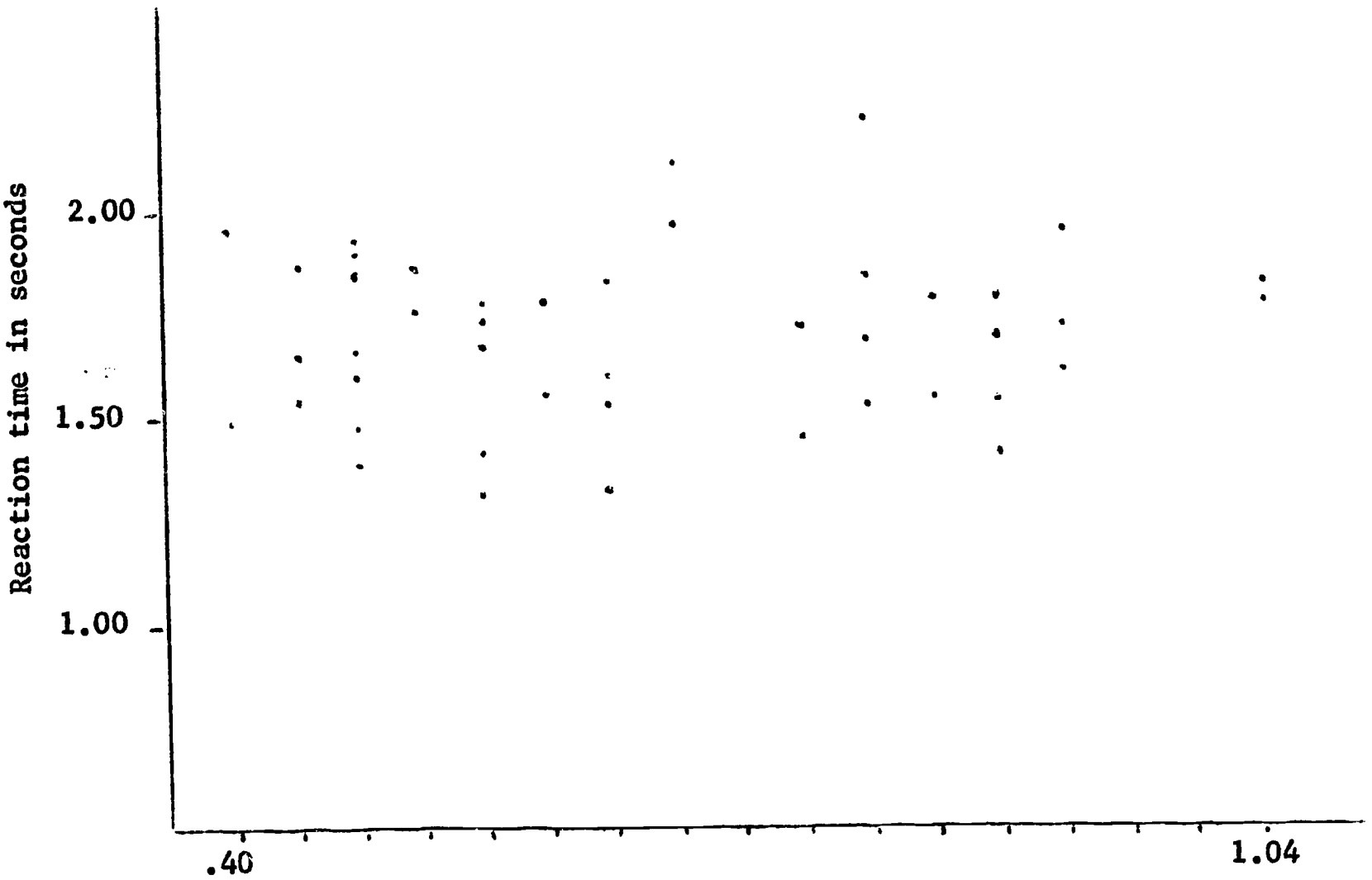

Figure 3. Scatterplot of the relationship between the length of the semantic experimental words and the latency of responses.